

별첨 사본은 아래 출원의 원본과 동일함을 증명함.

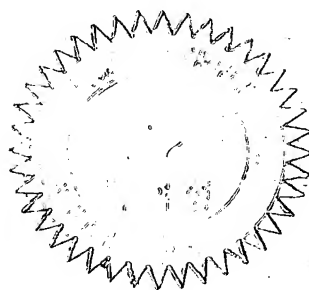
This is to certify that the following application annexed hereto is a true copy from the records of the Korean Intellectual Property Office.

출원 번호 : 10-2000-0082179  
Application Number

출원 년 월 일 : 2000년 12월 26일  
Date of Application DEC 26, 2000

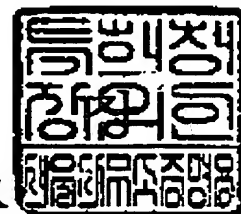
출원인 : 한국전자통신연구원  
Applicant(s) Electronics and Telecommunications Research Institute

2006년 04월 12일



특허청

COMMISSIONER



◆ This certificate was issued by Korean Intellectual Property Office. Please confirm any forgery or alteration of the contents by an issue number or a barcode of the document below through the KIPOnet- Online Issue of the Certificates' menu of Korean Intellectual Property Office homepage ([www.kipo.go.kr](http://www.kipo.go.kr)). But please notice that the

**【서지사항】**

<b>【서류명】</b>	명세서 등 보정서
<b>【수신처】</b>	특허청장
<b>【제출일자】</b>	2003.08.27
<b>【제출인】</b>	
<b>【명칭】</b>	한국전자통신연구원
<b>【출원인코드】</b>	3-1998-007763-8
<b>【사건과의 관계】</b>	출원인
<b>【대리인】</b>	
<b>【명칭】</b>	특허법인 신성
<b>【대리인코드】</b>	9-2000-100004-8
<b>【지정된변리사】</b>	변리사 정지원, 변리사 원석희, 변리사 박해천
<b>【포괄위임등록번호】</b>	2000-051975-8
<b>【사건의 표시】</b>	
<b>【출원번호】</b>	10-2000-0082179
<b>【출원일자】</b>	2000.12.26
<b>【심사청구일자】</b>	2000.12.26
<b>【발명의 명칭】</b>	자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치 및 그 방법
<b>【제출원인】</b>	
<b>【발송번호】</b>	9-5-2003-0199876-10
<b>【발송일자】</b>	2003.05.30
<b>【보정할 서류】</b>	명세서등
<b>【보정할 사항】</b>	
<b>【보정대상항목】</b>	별지와 같음
<b>【보정방법】</b>	별지와 같음
<b>【보정내용】</b>	별지와 같음

【취지】 특허법시행규칙 제13조 실용신안법시행규칙 제8조의 규정에 의하여 위  
와 같 이 제출합니다.

대리인

특허법인 신성 (인)

【수수료】

【보정료】 0 원

【추가심사청구료】 0 원

【기타 수수료】 0 원

【합계】 0 원

【첨부서류】 1.보정내용을 증명하는 서류[발명의상세한설명 보정]\_1통

## 【보정서】

【보정대상항목】 식별번호 22

【보정방법】 정정

【보정내용】

<22> 이처럼 본 발명에서는, 사람들이 가지고 있는 색상과 형태에 대한 퍼지 (fuzzy)한 대표값을 자연어와 함께 등록하여, 사용자가 친숙한 자연어로 질의어를 입력하면 그것을 분석한 후에, 검색하여 검색 결과를 제공한다. 즉, 종래의 색상과 형태를 직접 질의 방식으로 처리하던 방법에서 자연어 처리부를 한 번 더 거치게 함으로써, 사용자에게 더 친숙한 이미지 검색 인터페이스를 제공한다.

【보정대상항목】 식별번호 30

【보정방법】 정정

【보정내용】

<30> 그리고, 색상/형태 단어 인식부(14)는 상기 자연어 처리부(13)에서 나온 결과리스트를 입력받아 각 결과리스트 중에서 색상을 나타내는 단어와 형태를 나타내는 단어로 구별하여 인식하는 기능을 한다. 또한, 상기 구문 분석기에서 분석한 수식관계 및 기타 패턴을 활용하여 각각의 조건 등을 자동으로 인식하도록 한다.

【보정대상항목】 식별번호 34

【보정방법】 정정

## 【보정내용】

<34>

그리고, 형태 정보의 경우에도 기존 유형과 사용자 정의 유형으로 구별하여 기존 유형에서는 일반적으로 이미지 검색 시스템에서 유형별 검색에 자주 등장하는 모형을 일반화하여 각각의 어휘와 에지 정보를 함께 데이터베이스화하고, 사용자 정의 유형에서는 사용자가 검색하기를 원하는 에지 정보를 직접 입력(그리기 또는 파일 로딩)하도록 하여 어휘와 함께 저장한다.

이를 좀 더 상세히 살펴보면 다음과 같다.

일반적으로 형태 정보를 어휘(용어)와 매칭하는 것은 색상 정보의 경우만큼 용이하지는 않다. 그 이유는 형태가 형체를 가진 객체의 자세(Pose)와, 영상 내에서의 위치(Position)와 크기에 따라 다르기 때문이다. 즉, 예를 들어 자동차라는 어휘를 형태로 매칭할 때, 어느 위치에서 보는 자동차를 "자동차" 어휘와 매칭시킬 것인가하는 문제가 생긴다. 앞면에서 보는 자동차를 대표 자동차로 볼 것인가? 옆면에서 보는 자동차를 대표 자동차로 볼 것인가? 또한, 이러한 자동차가 디지털 영상에서 어느 한 모서리에 작은 크기로 표현되었는가 또는 영상 전체에 큰 크기로 표현되었는가에 따라 형태를 표현한 에지 임계치에 대한 데이터베이스의 저장 값이 다를 것이다.

이러한 문제를 해결하여 형태를 표현한 어휘와 형태 자체를 매칭시키는 방법은 다음과 같다. 객체를 보는 각도에 따른 형태의 변화에 대해서는 객체의 다양한 각도에서의 에지 분포를 하나의 객체에 해당하는 여러 개의 에지 성분의 임계치로

잡으면 해결이 가능하다. 예를 들어, 자동차의 경우에 자동차의 대표적인 위치에 따라 앞면, 측면, 후면 등에서 바라본 에지 성분들을 임계치 데이터베이스로 구축한 다음에 자동차라는 질의어가 들어왔을 때 자동차의 에지 임계치에서 앞면과 측면 및 후면 모두를 검색하여 어느 하나라도 검색 값이 높으면 앞면과 측면 및 후면 모두에 의해서 검색하여 그 결과를 출력하여 준다. 그리고, 자동차 종류에 따라서 다를 수도 있지만, 이 때는 승용차, 트럭, 버스 등과 같이 비슷한 종류의 자동차들끼리 모아 학습에 의한 방식으로 에지 성분을 저장한다.

그리고, 이미지 내에서의 위치와 크기에 따른 어휘와의 매칭은 물체가 중앙에 와야 한다는 조건으로 한정하면 쉽게 구현할 수 있다. 그러나, 이러한 방식은 조건을 벗어나는 데이터를 검색할 수 없어 검색 정확도가 떨어지기 때문에, 객체 위치 파악(object detection) 기술을 이용하여 객체 위치를 파악하여 객체 부분을 추출한 후에, 그 객체 부분에 해당하는 에지 성분으로 데이터베이스를 구축한다.

상기와 같은 형태 정보 매칭 기술에 대해서는 클립아트 프로토타입 시스템을 직접 구현하여 그 효용성을 확인하였다. 이러한 클립아트는 색상특징 표현보다는 형태특징 표현을 강조한 영상세트로서, 형태를 표현한 어휘를 통한 검색이 용이하다.

그리고, 검색 결과 출력부(18)는 입력받은 질의문을 분석하여 검색한 결과인 이미지 데이터를 사용자에게 출력하여 준다.

【보정대상항목】 식별번호 35

【보정방법】 삭제

【보정대상항목】 식별번호 40

【보정방법】 정정

【보정내용】

<40>           다음으로, 색상/형태 임계치 데이터베이스 검색부(17)에서는 [붉다]의 임계치 정보(색상 히스토그램 데이터)를 미리 구축해 둔 색상/형태 데이터베이스(15)를 검색하여, [아주]라는 수식어구를 처리하기 위해 앞에서 찾은 붉은 색상의 임계치에서 명도와 채도를 높인 수치로 전환해 준다(29).

## 【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0016
【제출일자】	2000. 12. 26
【발명의 국문명칭】	자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치 및 그 방법
【발명의 영문명칭】	Color and shape search method and apparatus of image data based on natural language with fuzzy concept
【출원인】	
【명칭】	한국전자통신연구원
【출원인코드】	3-1998-007763-8
【대리인】	
【성명】	특허법인 신성 정지원
【대리인코드】	9-2000-000292-3
【포괄위임등록번호】	2000-051975-8
【대리인】	
【성명】	특허법인 신성 원석희
【대리인코드】	9-1998-000444-1
【포괄위임등록번호】	2000-051975-8
【대리인】	
【성명】	특허법인 신성 박해천
【대리인코드】	9-1998-000223-4
【포괄위임등록번호】	2000-051975-8
【발명자】	
【성명의 국문표기】	박성희
【성명의 영문표기】	PARK, Sung Hee



**【주민등록번호】** 720114-1927713  
**【우편번호】** 302-172  
**【주소】** 대전광역시 서구 갈마2동 푸른아파트 1109호  
**【국적】** KR  
**【발명자】**  
**【성명의 국문표기】** 김현진  
**【성명의 영문표기】** KIM,Hyun Jin  
**【주민등록번호】** 710804-2094210  
**【우편번호】** 305-345  
**【주소】** 대전광역시 유성구 신성동 하나아파트 110-703  
**【국적】** KR  
**【발명자】**  
**【성명의 국문표기】** 박수준  
**【성명의 영문표기】** PARK,Soo Jun  
**【주민등록번호】** 660112-1052431  
**【우편번호】** 135-111  
**【주소】** 서울특별시 강남구 압구정1동 447 현대아파트 203-501  
**【국적】** KR  
**【발명자】**  
**【성명의 국문표기】** 장명길  
**【성명의 영문표기】** JANG,Myung Gil  
**【주민등록번호】** 650522-1095018  
**【우편번호】** 305-335  
**【주소】** 대전광역시 유성구 궁동 395-1 다솔아파트 101-102  
**【국적】** KR  
**【심사청구】** 청구

**【취지】** 특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출

원심사 를 청구합니다.

대리인	특허법인 신성 정지원 (인)
대리인	특허법인 신성 원석희 (인)
대리인	특허법인 신성 박해천 (인)

**【수수료】**

【기본출원료】 20 면 29,000 원

【가산출원료】 2 면 2,000 원

【우선권주장료】 0 건 0 원

【심사청구료】 8 항 365,000 원

【합계】 396,000 원

【감면사유】 정부출연연구기관

【감면후 수수료】 198,000 원

【첨부서류】 1. 요약서 명세서(도면)\_1통

## 【요약서】

### 【요약】

#### 1. 청구범위에 기재된 발명이 속한 기술분야

본 발명은 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치 및 그 방법에 관한 것임.

#### 2. 발명이 해결하려고 하는 기술적 과제

본 발명은 사람들이 가지고 있는 색상과 형태에 대한 퍼지(fuzzy)한 대표값을 자연어와 함께 등록하여, 사용자가 친숙한 자연어로 입력하면 질의어를 분석하여 검색하여 주는 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치 및 그 방법과 상기 방법을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공하는데 그 목적이 있음.

#### 3. 발명의 해결방법의 요지

본 발명은, 언어 처리용 사전 정보를 저장하고 있는 언어 처리용 사전 저장 수단; 색상과 관련된 단어와 매핑된 색상 히스토그램(histogram)값과, 형태(shape)와 관련된 단어에 상당하는 에지 정보를 저장하고 있는 색상/형태 임계치 저장 수단; 사용자로부터 색상과 형태를 기술한 자연어 질의 문장을 입력받기 위한 질의문 입력 수단; 상기 질의문 입력 수단을 통하여 입력된 자연어 문장을 상기 언어 처리용 사전 저장 수단에 저장되어 있는 사전 정보를 이용하여 분석하기 위한 자연어 처리 수단; 상기 자연어 처리 수단에서 분석된 단어에 대하여 색상 또는 형태를

나타내는 단어인지를 인식하기 위한 색상/형태 단어 인식 수단; 색상과 관련된 단어와 색상 히스토그램값을 매핑하여 상기 색상/형태 임계치 저장 수단에 저장하고, 형태와 관련된 단어에 상당하는 에지 정보를 상기 색상/형태 임계치 저장 수단에 저장하기 위한 색상/형태 임계치 데이터베이스 구축 수단; 상기 색상/형태 단어 인식 수단에서 인식된 색상과 형태에 대한 요구를 만족하는 데이터를 상기 색상/형태 임계치 저장 수단에서 검색하기 위한 색상/형태 임계치 검색 수단; 및 상기 색상/형태 임계치 검색 수단에서 검색한 이미지 데이터를 상기 사용자에게 제공하기 위한 검색 결과 출력 수단을 포함한다.

#### 4. 발명의 중요한 용도

본 발명은 이미지 데이터의 검색 등에 이용됨.

#### 【대표도】

도 1

#### 【색인어】

자연어, 색상, 형태, 검색, 이미지 데이터

【명세서】

### 【발명의 명칭】

자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치 및 그 방법{Color  
and shape search method and apparatus of image data based on natural language  
with fuzzy concept}

### 【도면의 간단한 설명】

- <1> 도 1 은 본 발명에 따른 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치의 일실시에 구성도.
- <2> 도 2 는 본 발명에 따른 자연어를 이용한 이미지 데이터 색상 및 형태 검색 방법에 대한 일실시에 흐름도.
- <3> 도 3 은 본 발명에 따른 데이터 흐름에 대한 일예시도.
- <4> \* 도면의 주요 부분에 대한 부호의 설명
- <5> 11 : 질의문 입력부                      12 : 언어 처리용 사전 저장부
- <6> 13 : 자연어 처리부                      14 : 색상/형태 단어 인식부
- <7> 15 : 색상/형태 임계치 데이터베이스
- <8> 16 : 색상/형태 임계치 데이터베이스 구축부
- <9> 17 : 색상/형태 임계치 데이터베이스 검색부

<10> 18 : 검색 결과 출력부

**【발명의 상세한 설명】**

**【발명의 목적】**

**【발명이 속하는 기술분야 및 그 분야의 종래기술】**

<11> 본 발명은 기존의 이미지 데이터의 검색 방식인 색상과 형태 정보 검색 방식과 자연어 질의 방식을 혼합한 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치 및 그 방법과 상기 방법을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체에 관한 것으로서, 더욱 상세하게는 사람들이 가지고 있는 색상과 형태에 대한 퍼지(fuzzy)한 대표값을 자연어와 함께 등록하여, 사용자가 친숙한 자연어로 입력하면 질의어를 분석하여 검색하여 주는 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치 및 그 방법과 상기 방법을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체에 관한 것이다.

<12> 현재 이미지 데이터를 검색하는 질의 방식에서는 색상과 형태 정보를 직접 입력하는 방식을 따르고 있다. 그래서, 사용자가 직접 검색을 원하는 색상에 대한 RGB(Red, Green, Blue)값을 숫자 형식으로 직접 입력하던지, 색상이 있는 팔레트에서 비슷한 색상을 고르는 질의 방식을 취하도록 하고 있다. 또한, 형태 정보(비슷한 에지 정보를 가진 이미지를 검색하는 방식)를 직접 이미지로 그리거나, 아니면 원하는 정보와 유사한 형태를 가진 그림을 직접 입력하는 질의 방식으로 검색하도록 하고 있다.

<13> 그러나, 상기와 같은 종래의 방법은, 데이터베이스로 확보해야 하는 어휘 정보(사전 정보)가 적다는 장점이 있으나, 이미지 데이터 검색 시스템의 입력으로 들어오는 질의문의 경우에 색상과 형태로 제한되어 있어서, 색상과 형태에 대한 각각의 수치들을 미리 대응시켜 두어야 하는 절차가 필요하다.

<14> 한편, 물론 몇몇 이미지 검색 시스템에서는 이러한 자연어 처리 기법을 응용하여 색상을 나타내는 어휘들을 이용한 질의 검색을 시도하고 있으나, 이는 시스템에서 지정한 한정된 어휘에 한해서만 가능하도록 하고 있다.

<15> 즉, 자연어로 검색하는 기술은 기존 정보검색에 많이 적용되어서 보편화되었지만, 이미지 정보검색은 주석정보(contents)를 바탕으로 한 검색에만 자연어 처리 기술을 적용하기 때문에 실제 칼라 검색이나 형태 검색의 경우에는 사용자가 친숙한 자연어로 된 검색이 불가능한 문제점이 있었다.

#### 【발명이 이루고자 하는 기술적 과제】

<16> 본 발명은 상기 문제점을 해결하기 위하여 제안된 것으로, 사람들이 가지고 있는 색상과 형태에 대한 퍼지(fuzzy)한 대표값을 자연어와 함께 등록하여, 사용자가 친숙한 자연어로 입력하면 질의어를 분석하여 검색하여 주는 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치 및 그 방법과 상기 방법을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공하는데 그 목적이 있다.

## 【발명의 구성】

<17>

상기 목적을 달성하기 위한 본 발명의 장치는, 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치에 있어서, 언어 처리용 사전 정보를 저장하고 있는 언어 처리용 사전 저장 수단; 색상과 관련된 단어와 매핑된 색상 히스토그램(histogram)값과, 형태(shape)와 관련된 단어에 상당하는 에지 정보를 저장하고 있는 색상/형태 임계치 저장 수단; 사용자로부터 색상과 형태를 기술한 자연어 질의 문장을 입력받기 위한 질의문 입력 수단; 상기 질의문 입력 수단을 통하여 입력된 자연어 문장을 상기 언어 처리용 사전 저장 수단에 저장되어 있는 사전 정보를 이용하여 분석하기 위한 자연어 처리 수단; 상기 자연어 처리 수단에서 분석된 단어에 대하여 색상 또는 형태를 나타내는 단어인지를 인식하기 위한 색상/형태 단어 인식 수단; 색상과 관련된 단어와 색상 히스토그램값을 매핑하여 상기 색상/형태 임계치 저장 수단에 저장하고, 형태와 관련된 단어에 상당하는 에지 정보를 상기 색상/형태 임계치 저장 수단에 저장하기 위한 색상/형태 임계치 데이터베이스 구축 수단; 상기 색상/형태 단어 인식 수단에서 인식된 색상과 형태에 대한 요구를 만족하는 데이터를 상기 색상/형태 임계치 저장 수단에서 검색하기 위한 색상/형태 임계치 검색 수단; 및 상기 색상/형태 임계치 검색 수단에서 검색한 이미지 데이터를 상기 사용자에게 제공하기 위한 검색 결과 출력 수단을 포함하는 것을 특징으로 한다.

<18>

한편, 본 발명의 방법은, 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치에 적용되는 이미지 데이터 색상 및 형태 검색 방법에 있어서, 언어 처리용 사



전 정보를 저장하고, 색상과 관련된 단어와 매핑된 색상 히스토그램(histogram)값과, 형태(shape)와 관련된 단어에 상당하는 에지 정보를 저장하는 제 1 단계; 사용자로부터 입력된 색상과 형태를 기술한 자연어 질의 문장을 상기 사전 정보를 이용하여 분석하는 제 2 단계; 상기 분석된 단어에 대하여 색상 또는 형태를 나타내는 단어인지를 인식하는 제 3 단계; 상기 인식된 색상과 형태에 대한 요구를 만족하는 데이터를 상기 저장된 색상 히스토그램(histogram)값과 에지 정보에서 검색하는 제 4 단계; 및 상기 검색한 이미지 데이터를 상기 사용자에게 제공하는 제 5 단계를 포함하는 것을 특징으로 한다.

<19> 또한, 상기 본 발명의 방법은, 상기 검색시 상기 인식된 색상과 형태에 대한 요구를 만족하는 데이터가 없을 경우에, 상기 사용자로부터 상기 인식된 색상과 형태에 대한 임계치를 입력받아 저장하는 제 6 단계를 더 포함하는 것을 특징으로 한다.

<20> 한편, 본 발명은, 자연어를 이용하여 이미지 데이터 색상 및 형태를 검색하기 위하여, 프로세서를 구비한 이미지 데이터 색상 및 형태 검색 장치에, 언어 처리용 사전 정보를 저장하고, 색상과 관련된 단어와 매핑된 색상 히스토그램(histogram)값과, 형태(shape)와 관련된 단어에 상당하는 에지 정보를 저장하는 제 1 기능; 사용자로부터 입력된 색상과 형태를 기술한 자연어 질의 문장을 상기 사전 정보를 이용하여 분석하는 제 2 기능; 상기 분석된 단어에 대하여 색상 또는 형태를 나타내는 단어인지를 인식하는 제 3 기능; 상기 인식된 색상과 형태에 대한 요구를 만족하는 데이터를 상기 저장된 색상 히스토그램(histogram)값과 에지 정보에

서 검색하는 제 4 기능; 및 상기 검색한 이미지 데이터를 상기 사용자에게 제공하는 제 5 기능을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공한다.

<21> 또한, 상기 본 발명은, 상기 검색시 상기 인식된 색상과 형태에 대한 요구를 만족하는 데이터가 없을 경우에, 상기 사용자로부터 상기 인식된 색상과 형태에 대한 임계치를 입력받아 저장하는 제 6 기능을 더 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공한다.

<22> 이처럼 본 발명에서는, 사람들이 가지고 있는 색상과 형태에 대한 퍼지(fuzzy)한 대표값을 자연어와 함께 등록하여, 사용자가 친숙한 자연어로 질의어를 입력하면 그것을 분석하여, 검색하여 검색 결과를 제공한다. 즉, 종래의 색상과 형태를 직접 질의 방식으로 처리하던 방법에서 자연어 처리부를 한 번 더 거치게 함으로써, 사용자에게 더 친숙한 이미지 검색 인터페이스를 제공한다.

<23> 또한, 본 발명에서는 자연어로 입력되는 단순한 명사 형태의 질의어 뿐만 아니라, 색상과 형태를 복합으로 처리할 수 있도록 하기 위해, 수식문을 포함한 단문 형태로 확장할 수 있도록 한다.

<24> 예를 들어, "붉은 색의 산"이라는 문장을 처리하기 위해서는 "붉다"라는 단어에 대해서 먼저 수식어라는 단어의 구문적인 정보 추출과 색상을 나타내는 단어라는 것을 인식하는 자연어 처리와 함께 붉은 색을 나타내는 RGB값의 임계치를 데이터베이스로부터 가져와야 한다. 또한, "산"이라는 단어에 대해서는 형태 정보로 인식하고, "산"에 대한 예지 임계치를 색상과 같이 데이터베이스로부터 가져와서

검색을 원하는 색상과 형태를 가진 이미지 데이터를 검색한다.

<25> 상술한 목적, 특징들 및 장점은 첨부된 도면과 관련한 다음의 상세한 설명을 통하여 보다 분명해 질 것이다. 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 일실시예를 상세히 설명한다.

<26> 도 1 은 본 발명에 따른 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치의 일실시예 구성도이다.

<27> 도 1 에 도시된 바와 같이, 본 발명에 따른 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치는, 사용자가 검색하기를 원하는 색상과 형태를 기술한 자연어 질의 문장을 입력받기 위한 질의문 입력부(11), 언어 처리용 사전 정보를 저장하고 있는 언어 처리용 사전 저장부(12), 상기 질의문 입력부(11)를 통하여 입력된 자연어를 상기 언어 처리용 사전 저장부(12)에 저장되어 있는 사전 정보를 이용하여 분석하기 위한 자연어 처리부(13), 상기 자연어 처리부(13)에서 분석된 단어에 대하여 색상 또는 형태를 나타내는 단어인지를 인식하고 각각의 조건을 인식하는 색상/형태 단어 인식부(14), 색상과 관련된 단어와 매핑된 색상 히스토그램(histogram)값과, 형태(shape)와 관련된 단어에 해당하는 에지 정보를 저장하고 있는 색상/형태 임계치 데이터베이스(15), 색상과 관련된 단어와 색상 히스토그램(histogram)값을 매핑(mapping)하여 상기 색상/형태 임계치 데이터베이스(15)에 저장하고, 형태(shape)와 관련된 단어에 해당하는 에지 정보를 상기 색상/형태 임계치 데이터베이스(15)에 저장하기 위한 색상/형태 임계치 데이터베이스 구축부(16), 상기 색상/형태 단어 인식부(14)에서 인식된 색상과 형태에 대한 요구를 만족하는

데이터를 상기 색상/형태 임계치 데이터베이스(15)에서 검색하기 위한 색상/형태 임계치 데이터베이스 검색부(17) 및 색상/형태 임계치 데이터베이스 검색부(17)에서 검색한 이미지 데이터를 상기 사용자에게 제공하기 위한 검색 결과 출력부(18)를 포함한다.

<28> 이제, 본 발명에 따른 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치의 중요 구성요소에 대하여 좀 더 상세히 살펴보면 다음과 같다.

<29> 먼저, 자연어 처리부(13)는 사용자가 검색하기를 원하는 색상과 형태를 기술한 질의 문장을 질의문 입력부(11)를 통하여 입력받아 자연어 분석을 하는 모듈로서, 입력 문장에 대하여 단어를 분리 분석하는 부분이다. 따라서, 자연어 처리부(13)는 크게 입력문을 어절 단위로 분리하는 어절 분리기(Tokenizer)와 어절 분리기에서 나온 출력 데이터에서 각 단어의 사전 정보를 검색하고 형태소적 정보를 찾아 내는 사전 검색기와 구문 형식을 파악하는 구문 분석기(Parser)를 포함한다.

<30> 그리고, 색상/형태 단어 인식부(14)는 상기 자연어 처리부(13)에서 나온 결과 리스트를 입력받아 각 결과리스트 중에서 색상을 나타내는 단어와 형태를 나타내는 단어로 구별하여 인식하는 기능을 하고 있다. 또한, 상기 구문 분석기에서 분석한 수식관계 및 기타 패턴을 활용하여 각각의 조건 등을 자동으로 인식하도록 한다.

<31> 그리고, 색상/형태 임계치 데이터베이스 구축 및 검색부(16,17)는 색상/형태 단어 인식부(14)에서 인식된 결과로 나온 색상과 형태에 대한 요구를 만족하는 데이터에 대해서 검색하고, 미리 이러한 색상과 형태에 대한 데이터를 저장하는 기능

을 수행한다. 즉, 미리 사용자가 요구할 만한 정보에 대해서 각각의 임계치값을 저장하여 데이터베이스를 구축하고, 검색부를 통해서 검색한다.

<32> 이때, 색상(color) 질의에서는 기존 색상을 나타내는 단어와 색상 히스토그램(histogram)값을 매핑(mapping)하여 데이터베이스화하는 방식(예 : 초록, 푸른색, 붉은색 등)과 색상을 나타내지는 않더라도 개념적으로 색을 표현할 수 있는 어휘 또는 개념과 그 단어에 맞는 색상 히스토그램(histogram)값을 함께 저장하는 방식(예 : 단풍, 바다, 노을 등)을 사용한다. 또한, 형태(shape) 질의에서는 객관화될 수 있는 형태를 가진 어휘에 상응하는 에지 정보를 데이터베이스화하여 저장하는 방식을 사용한다.

<33> 그리고, 상기 색상 정보의 경우에는 기존 팔레트와 사용자 정의 팔레트 정보로 구별하여 기존 팔레트 정보를 이용하는 경우에는 256가지 색상에 대해서 각각의 어휘와 임계치를 매핑(Mapping)하여 데이터베이스로 구축하고, 사용자 팔레트 정보를 사용하는 경우에는 사용자가 직접 팔레트 판을 이용하여 색상을 지정한 후 거기에 해당하는 어휘를 입력받아 저장하여, 이후에는 그 어휘에 대해서도 검색이 자동으로 계속되도록 하는 기능을 제공한다.

<34> 그리고, 형태 정보의 경우에도 기존 유형과 사용자 정의 유형으로 구별하여 기존 유형에서는 일반적으로 이미지 검색 시스템에서 유형별 검색에 자주 등장하는 모형을 일반화하여 각각의 어휘와 에지 정보를 함께 데이터베이스화하고, 사용자 정의 유형에서는 사용자가 검색하기를 원하는 에지 정보를 직접 입력(그리기 또는 파일 로딩)하도록 하여 어휘와 함께 저장한다.

- <35> 그리고, 검색 결과 출력부(18)는 입력받은 질의문을 분석하여 검색한 결과인 이미지 데이터를 사용자에게 출력하여 준다.
- <36> 도 2 는 본 발명에 따른 자연어를 이용한 이미지 데이터 색상 및 형태 검색 방법에 대한 일실시에 흐름도로서, 사용자가 "아주 붉은 색"이라는 질의문을 입력한 경우를 예로 들어 설명하기로 한다.
- <37> 먼저, 사용자가 "아주 붉은 색"과 같이 자연어로 된 질의문을 입력하면, 자연어 처리부(13)의 어휘 분석기에서는 [아주] [붉은] [색]의 세 어휘로 분리한다(21). 그러면, 사전 검색기가 언어 처리용 사전 저장부(12)를 통해서 [아주-(부사)] [붉-(형용사)+은-(관형형 어미)] [색-(색상을 통칭하는 명사)]을 검색하여 n 개(n은 자연수)의 어절 후보 리스트를 생성한다(22). 이후, 상기 n개의 어절 후보에 대하여 색상/형태 정보를 검색하고 구문 관계를 설정하는 과정을 반복 수행한다(23 내지 27). 이때, 사전 검색기로부터 어절 후보 리스트를 입력으로 받은 구문 분석기에서는 상기 입력 문장에서 [아주]라는 부사가 [붉은]을 수식하고 [아주 붉은]이 [색]을 수식하는 상관 관계가 있음을 분석해 낸다.
- <38> 상기와 같이 분석된 결과를 입력받은 색상/형태 단어 인식부(14)에서는 형용사와 명사 정보를 위주로 상기 단어들 중에 색상 또는 형태를 나타내는 단어를 인식한다(28).
- <39> 이처럼, 형용사로 분석된 [붉다]의 경우와 [색]이라는 명사를 사전 정보를 통해서 색상을 나타내는 형용사와 색상을 일컫는 통칭이라는 정보로 인식해 내고, [아주]라는 부사는 색상의 경우에는 바로 뒤의 피수식어구를 강조하는 의미가 있음

을 인식한다.

<40>           다음으로, 색상/형태 임계치 데이터베이스 검색부(17)에서는 [붉다]의 임계치 정보(색상 히스토그램 데이터)를 미리 구축해 둔 색상/형태 데이터베이스(15)를 검색하여 찾아, [아주]라는 수식어구를 처리하기 위해 앞에서 찾은 붉은 색상의 임계치에서 명도와 채도를 높인 수치로 전환해 준다(29).

<41>           이후, 검색 결과 출력부(18)에서는 검색한 이미지 데이터를 사용자에게 출력한다(30).

<42>           한편, 색상/형태 데이터베이스(15)를 검색한 결과, 찾고자 하는 색상 임계치 정보(색상 히스토그램 데이터)가 없으면 색상/형태 데이터를 사용자가 등록한 후에 (31) 검색하도록 한다.

<43>           도 3 은 본 발명에 따른 데이터 흐름에 대한 일예시도이다.

<44>           도 3 에 도시된 바와 같이, "초록색 배경에 붉은 색 자동차가 있는 그림"이라는 문장이 입력되면, 어절 후보 리스트로는 {초록색}, {배경+에}, {붉+은}, {색}, {자동차+가}, {있+는}, {그림}이 생성된다.

<45>           그리고, 어절별 색상/형태 정보 검색 과정에서는 {초록색} : 초록(green)(색상(color)), {배경+에} : 배경(background)(팩터(factor)), {붉+은}{색} : 레드(red)(색상(color)), {자동차+가} : 자동차(car)(형태(shape)), {있+는} : 없음(none), {그림} : 그림(image)(팩터(factor)) 등으로 검색된다.

<46>           그리고, 구문 관계 설정 과정에서는 "초록색"이 "배경에"를 수식하고,

"붉은"이 "색"을 수식하고, "붉은 색"이 "자동차가"를 수식하고, "초록색 배경에"와 "붉은 색 자동차가"가 "있는"을 수식하고, "초록색 배경에 붉은 색 자동차가 있는"이 "그림"을 수식한다.

<47> 그리고, 색상/형태 데이터베이스 검색 과정에서는 "초록색"과 "배경"을 색상 데이터베이스에서 검색하고, "붉은색"과 "자동차"를 색상 데이터베이스와 형태 데이터베이스에서 검색한다.

<48> 이후, 검색 결과를 사용자에게 제시한다.

<49> 상술한 바와 같은 본 발명의 방법은 프로그램으로 구현되어 컴퓨터로 읽을 수 있는 형태로 기록매체(씨디롬, 램, 롬, 플로피 디스크, 하드 디스크, 광자기 디스크 등)에 저장될 수 있다.

<50> 이상에서 설명한 본 발명은 전술한 실시예 및 첨부된 도면에 의해 한정되는 것이 아니고, 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 여러 가지 치환, 변형 및 변경이 가능하다는 것이 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 있어 명백할 것이다.

### 【발명의 효과】

<51> 상기와 같은 본 발명은, 사람들이 가지고 있는 색상과 형태에 대한 퍼지(fuzzy)한 대표값을 자연어와 함께 등록하여, 사용자가 친숙한 자연어로 질의어를 입력하면 질의어를 분석하여 검색하여 줌으로써, 기존의 색상 배합과 단순한 형태



정보만을 이용한 질의 방식을 개선시킬 수 있는 효과가 있다.

<52>           이처럼, 본 발명은 색상(color)과 형태(shape) 검색에 자연어 처리 기능을 부가함으로써, 사용자에게 친숙한 검색 인터페이스를 제공하여 사용자의 편리성을 크게 향상시킬 수 있는 효과가 있다.

## 【특허청구범위】

### 【청구항 1】

자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치에 있어서,

언어 처리용 사전 정보를 저장하고 있는 언어 처리용 사전 저장 수단;

색상과 관련된 단어와 매핑된 색상 히스토그램(histogram)값과, 형태(shape)와 관련된 단어에 상당하는 에지 정보를 저장하고 있는 색상/형태 임계치 저장 수단;

사용자로부터 색상과 형태를 기술한 자연어 질의 문장을 입력받기 위한 질의 문 입력 수단;

상기 질의문 입력 수단을 통하여 입력된 자연어 문장을 상기 언어 처리용 사전 저장 수단에 저장되어 있는 사전 정보를 이용하여 분석하기 위한 자연어 처리 수단;

상기 자연어 처리 수단에서 분석된 단어에 대하여 색상 또는 형태를 나타내는 단어인지를 인식하기 위한 색상/형태 단어 인식 수단;

색상과 관련된 단어와 색상 히스토그램값을 매핑하여 상기 색상/형태 임계치 저장 수단에 저장하고, 형태와 관련된 단어에 상당하는 에지 정보를 상기 색상/형태 임계치 저장 수단에 저장하기 위한 색상/형태 임계치 데이터베이스 구축 수단;

상기 색상/형태 단어 인식 수단에서 인식된 색상과 형태에 대한 요구를 만족하는 데이터를 상기 색상/형태 임계치 저장 수단에서 검색하기 위한 색상/형태 임

계치 검색 수단; 및

상기 색상/형태 임계치 검색 수단에서 검색한 이미지 데이터를 상기 사용자에게 제공하기 위한 검색 결과 출력 수단을

을 포함하는 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치.

### 【청구항 2】

제 1 항에 있어서,

상기 색상/형태 임계치 검색 수단에서 검색시 상기 색상/형태 단어 인식 수단에서 인식된 색상과 형태에 대한 요구를 만족하는 데이터가 없을 경우에, 상기 색상/형태 임계치 데이터베이스 구축 수단이 상기 사용자로부터 상기 인식된 색상과 형태에 대한 임계치를 입력받아 상기 색상/형태 임계치 저장 수단에 저장하는 것을 특징으로 하는 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치.

### 【청구항 3】

제 1 항 또는 제 2 항에 있어서,

상기 색상/형태 임계치 데이터베이스 구축 수단은,

색상(color) 질의어에 대하여 기존 색상을 나타내는 단어와 색상 히스토그램(histogram)값을 매핑(mapping)하여 데이터베이스화하고, 색상을 나타내지는 않더라도 개념적으로 색을 표현할 수 있는 어휘 또는 개념과 그 단어에 맞는 색상 히스

토그램(histogram)값을 함께 저장하며, 형태(shape) 질의어에 대하여 객관화될 수 있는 형태를 가진 어휘에 상당하는 에지 정보를 데이터베이스화하여 저장하는 것을 특징으로 하는 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치.

#### 【청구항 4】

제 3 항에 있어서,

상기 색상/형태 단어 인식 수단은,

상기 자연어 처리 수단에서 분석한 수식관계 및 기타 패턴을 활용하여 각각의 조건을 자동으로 인식하는 것을 특징으로 하는 자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치.

#### 【청구항 5】

자연어를 이용한 이미지 데이터 색상 및 형태 검색 장치에 적용되는 이미지 데이터 색상 및 형태 검색 방법에 있어서,

언어 처리용 사전 정보를 저장하고, 색상과 관련된 단어와 매핑된 색상 히스토그램(histogram)값과, 형태(shape)와 관련된 단어에 상당하는 에지 정보를 저장하는 제 1 단계;

사용자로부터 입력된 색상과 형태를 기술한 자연어 질의 문장을 상기 사전 정보를 이용하여 분석하는 제 2 단계;

상기 분석된 단어에 대하여 색상 또는 형태를 나타내는 단어인지를 인식하는 제 3 단계;

상기 인식된 색상과 형태에 대한 요구를 만족하는 데이터를 상기 저장된 색상 히스토그램(histogram)값과 에지 정보에서 검색하는 제 4 단계; 및

상기 검색한 이미지 데이터를 상기 사용자에게 제공하는 제 5 단계

를 포함하는 자연어를 이용한 이미지 데이터 색상 및 형태 검색 방법.

#### 【청구항 6】

제 5 항에 있어서,

상기 검색시 상기 인식된 색상과 형태에 대한 요구를 만족하는 데이터가 없을 경우에, 상기 사용자로부터 상기 인식된 색상과 형태에 대한 임계치를 입력받아 저장하는 제 6 단계

를 더 포함하는 자연어를 이용한 이미지 데이터 색상 및 형태 검색 방법.

#### 【청구항 7】

자연어를 이용하여 이미지 데이터 색상 및 형태를 검색하기 위하여, 프로세서를 구비한 이미지 데이터 색상 및 형태 검색 장치에,

언어 처리용 사전 정보를 저장하고, 색상과 관련된 단어와 매핑된 색상 히스토그램(histogram)값과, 형태(shape)와 관련된 단어에 상당하는 에지 정보를 저장

하는 제 1 기능;

사용자로부터 입력된 색상과 형태를 기술한 자연어 질의 문장을 상기 사전 정보를 이용하여 분석하는 제 2 기능;

상기 분석된 단어에 대하여 색상 또는 형태를 나타내는 단어인지를 인식하는 제 3 기능;

상기 인식된 색상과 형태에 대한 요구를 만족하는 데이터를 상기 저장된 색상 히스토그램(histogram)값과 에지 정보에서 검색하는 제 4 기능; 및

상기 검색한 이미지 데이터를 상기 사용자에게 제공하는 제 5 기능

을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

#### 【청구항 8】

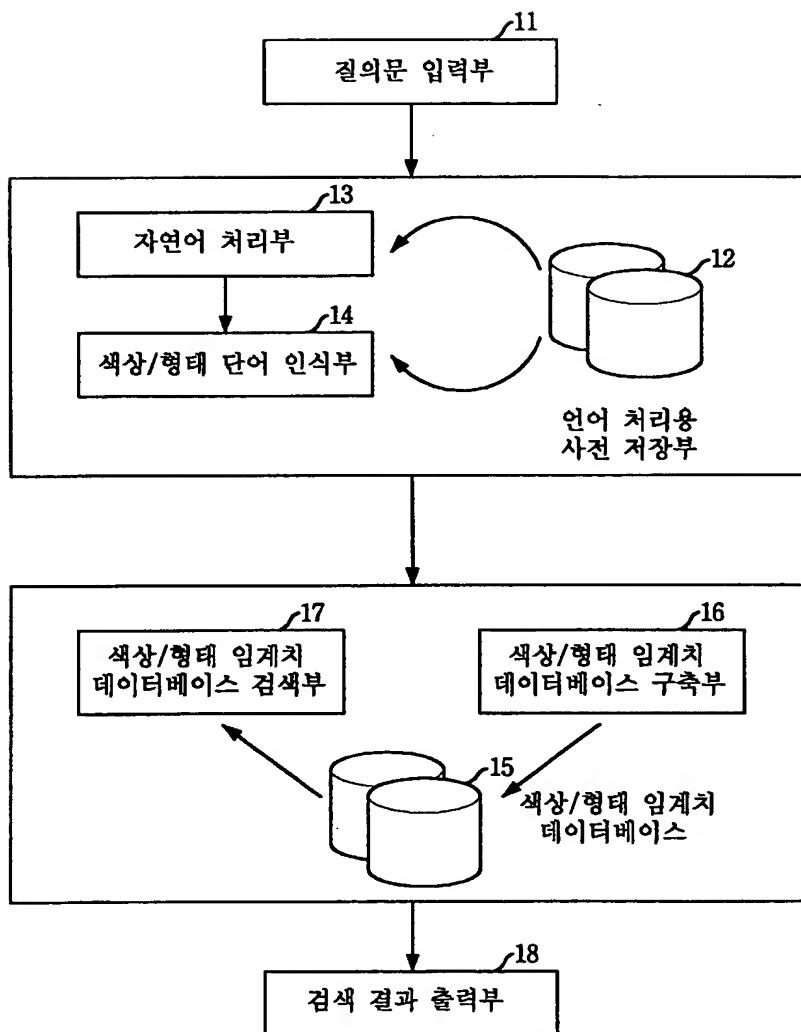
제 7 항에 있어서,

상기 검색시 상기 인식된 색상과 형태에 대한 요구를 만족하는 데이터가 없을 경우에, 상기 사용자로부터 상기 인식된 색상과 형태에 대한 임계치를 입력받아 저장하는 제 6 기능

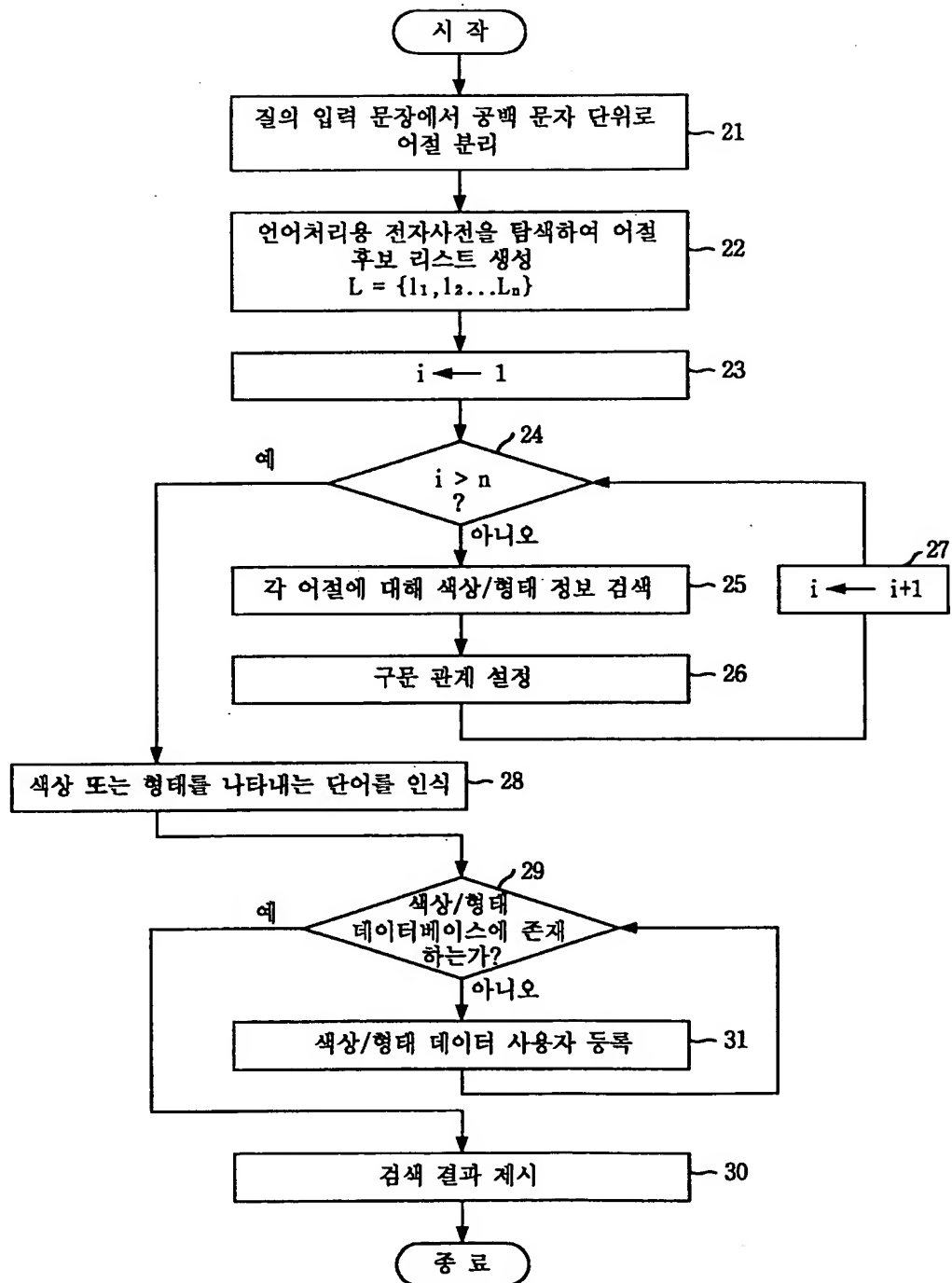
을 더 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

【도면】

【도 1】



【도 2】





【도 3】

